the generalized numerical range of matrix polynomials

Hiroshi Nakazato (Hirosaki University, JAPAN)
Abstract In this talk, the generalized numerical range of matrix polynomial is introduced and its properties are discussed. Especially the algorithm to compute the boundary of the generalized numerical range is provided in some cases.

Contents

00. Introduction  0. Motivations or Aims
1. Numerical objects of matrices and matrix polynomials
2. Numerical range of a matrix as $V^1(A)$
3. When I was younger,...
4. Weighted shift operators
5. (1) $C$-numerical range (2) application to NMR
6. Kippenhahn’s method to compute the boundary
7. (1) Mathematics and Philosophy (2) Al-Tusi (3) Epicycles and the $k$-numerical range
8. Joint numerical range and its convex hull
9. Polynomial numerical hulls
10. Conclusion

Fig.1 Lion stone carving at Persepolis
I would like to express my thanks to the organizers of this conference for inviting me and giving me an opportunity to talk here.

At first I call that Iran and Japan have a long history of their friendship. About 1600 years ago, the first Dynasty or State in Japan was built at Nara.

The photo in the next page is our national treasure which was brought from the Sasanian Dynasty via China about 1250 years ago.
Fig. 2  a National Treasure of Japan
That is an excellent cut glass.

At present an Iranian young athlete Ali Darvish is one of the best players of Japanese professional sports.

Fig.3  Ali Darvish
I hope that this visit promote our study on Mathematics and increase our friendship.

Fig. 4  a map of Asia
Fig.5  Mt. Fuji in Japan
0. Motivations or Aims

Matrix Analysis has many applications.

1: the application of the eigenvalue problem
to the analysis of the oscillations of buildings.

2: the applications to the analysis of
simultaneous differential equations.

3: Localization of the eigenvalues of a linear operator
In Iran and Japan, earth-quakes frequently occur. We are interested in the structural performance of buildings against earthquakes.

Fig.6 building destroyed by an earth-quake

In [19;2005], my colleague Tsumura and I discussed the strength of concrete columns against horizontal loads using numerical ranges. I hope that our study in Matrix Analysis contributes to the safety and economic activities of nations.
1. Numerical objects of matrices and matrix polynomials

Let $\mathbf{M}_n$ be the associative algebra of all $n \times n$ complex matrices. Suppose that

$$Q(\lambda) = A_m \lambda^m + A_{m-1} \lambda^{m-1} + \cdots + A_1 \lambda + A_0$$

is a matrix polynomial, where $A_i \in \mathbf{M}_n$ ($i = 0, 1, \ldots, m$) and $\lambda$ is a complex variable. Many numerical objects are defined for a matrix $A \in \mathbf{M}_n$ or a matrix polynomial $Q(\lambda)$. Probably, the most important one is their eigenvalues or the spectra. The spectrum $\sigma(A)$ of a matrix $A$ is defined as

$$\sigma(A) = \{\lambda \in \mathbb{C} : A\xi = \lambda \xi \text{ for some } \xi \in \mathbb{C}^n, \xi \neq 0\}$$

$$= \{\lambda \in \mathbb{C} : \det(\lambda I_n - A) = 0\}. \quad (1.1)$$
The spectra of matrices have good properties such as \( \sigma(p(A)) = p(\sigma(A)) \) for a polynomial \( p(\lambda) \) and \( \sigma(AB) = \sigma(BA) \). Let \( A = U|A| \) be the polar decomposition of \( A \). The Schatten \( p \)-norm \( ||A||_p \) of a matrix \( A \) is defined as

![Fig. 7  p=4](image)
\[ ||A||_p = \text{tr}(|A|^p)^{1/p} \]  
\[ (1 \leq p < \infty). \]

We also consider the operator norm \( ||A|| \) defined as

\[ ||A|| = \max\{ \sqrt{\langle A\xi\rangle^* \langle A\xi\rangle} : \xi \in \mathbb{C}^n, \xi^*\xi = 1 \} \]
\[ = \max\{ \sqrt{\lambda} : \lambda \geq 0, \det(\lambda I_n - A^*A) = 0 \}. \]  
\[ (1.3) \]
We consider the case $A = \text{diag}(a_1, a_2)$ with $a_1, a_2$ are real numbers. We consider the convex set

$$
\Omega = \{(a_1, a_2) \in \mathbb{R}^2 : \|\text{diag}(a_1, a_2)\|_p \leq 1\}
$$

$$
= \{(a_1, a_2) \in \mathbb{R}^2 : |a_1|^p + |a_2|^p \leq 1\}
$$

and its boundary $\partial \Omega$ in $\mathbb{R}^2$.

The curve $\partial \Omega$ is an algebraic curve, i.e., there is a real polynomial $f(x, y) \in \mathbb{R}[x, y]$ s.t. $f(a_1, a_2) = 0$ for $|a_1|^p + |a_2|^p = 1 \iff p$ is a rational number.

The spectrum $\sigma(Q(\lambda))$ of a matrix polynomial $Q(\lambda)$ is defined as

$$
\sigma(Q(\lambda)) = \{\lambda \in \mathbb{C} : Q(\lambda)\xi = 0, \text{ for some } \xi \in \mathbb{C}^n, \xi \neq 0\}
$$

$$
= \{\lambda \in \mathbb{C} : \det(Q(\lambda)) = 0\} = \{\lambda \in \mathbb{C} : 0 \in \sigma(Q(\lambda))\}. \quad (1.4)
$$
2. Numerical range of a matrix as $V^1(A)$

The numerical range $W(A)$ of a matrix $A \in M_n$ is defined as

$$W(A) = \{\xi^* A \xi : \xi \in \mathbb{C}^n, \xi^* \xi = 1\}.$$  \hfill (2.1)

In 1919, a German Mathematician Hausdorff [12] proved the convexity of $W(A)$. If $\lambda \in \mathbb{C}$ does not belong to $W(A)$, then the separation theorem implies that there exist $0 \leq \theta < 2\pi$ and $b \in \mathbb{R}$ satisfying

$$\Re(\exp(-i\theta)\xi^* A \xi) \leq b < \Re(\exp(-i\theta)\lambda).$$

for every $\xi \in \mathbb{C}^n$, $\xi^* \xi = 1$. It follows that the inequality

$$\xi^* [I_n + \frac{1}{a} (\exp(-i\theta)A + \exp(i\theta)A^*) + \frac{1}{a^2} A^* A] \xi$$

$$< 1 + \frac{2}{a} \Re(\exp(-i\theta)\lambda) + \frac{1}{a^2} |\lambda|^2$$
\((\xi \in \mathbb{C}^n, \xi^*\xi = 1)\) holds for sufficiently large \(a > 0\). Hence the inequality
\[||a \exp(-i\theta)I_n + A|| < |a \exp(-i\theta) + \lambda|\]
holds for sufficiently large \(a > 0\).

On the other hand, if the equation \(\lambda = \xi^*A\xi\) holds for some unit vector \(\xi\), then the inequality
\[|\lambda + \tilde{a}| \leq ||A + \tilde{a}I_n||\]
holds for \(\forall \tilde{a} \in \mathbb{C}\). Thus we have one characterization of \(W(A)\) as the following:

\[W(A) = V^1(A) = \{\lambda \in \mathbb{C} : |\lambda + a| \leq ||A + aI_n|| \text{ for } \forall a \in \mathbb{C}\}. \quad (2.2)\]
Fig. 8  a meeting of the numerical group-2006-16
3. When I was younger,…

The one-parameter semi-group \( \{ \exp(tA) : t \geq 0 \} \) generated by \( A \) is contractive, that is, \( \| \exp(tA) \| \leq 1 \) \( (t \geq 0) \) if and only if the condition

\[
W(A) \subset \{ \lambda \in \mathbb{C} : \Re(\lambda) \leq 0 \}
\]

holds.

About 25 years ago, I was studying such semi-groups on infinite dimensional spaces. I was also interested in one-parameter groups and their generator. It is my pleasure that one of my early papers [18;1984] in this subject was reviewed by Professor Asadollah Niknam in Mathematical Reviews. He was and is producing nice results in the theory of operator algebras (cf. [21;1982], [22;2000]). My main interests moved to Matrix Analysis about 15 years ago.
4. weighted shift operators

The numerical range is also defined for an operator in complex Hilbert space. Recently Chien and I found the following result.

**Theorem** (Chien-N; 2008) Let $T$ be a weighted shift operator

$$T = T(q) = \begin{pmatrix} 0 & 0 & 0 & \ldots \\ 1 & 0 & 0 & \ldots \\ 0 & q & 0 & \ldots \\ 0 & 0 & q^2 & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

with a geometric sequence $\{q^{n-1}\}_{n=1}^{\infty}$ as the weights ($0 < q < 1$). Then the numerical range $W(T)$ is the closed circular disc with center 0 with radius $w(T(q))$, where $1/(2w(T))$ is the minimum positive root of
a $q$-hypergeometric function

$$F(z) = 1 + \sum_{n=1}^{\infty} \frac{(-1)^n q^{2n(n-1)}}{(1 - q^2)(1 - q^4)(1 - q^6) \cdots (1 - q^{2n})} z^{2n}.$$ 

Fig. 9 the numerical radius $w(T(q))$
5.1 \( C \)-numerical range

Let \( C \) be a real diagonal matrix \( \text{diag}(c_1, c_2, \ldots, c_n) \). Then the range \( W_C(A) \) is defined for \( A \in M_n \) as

\[
W_C(A) = \left\{ \sum_{j=1}^{n} c_j \xi_j^* A \xi_j : \{\xi_1, \xi_2, \ldots, \xi_n\} \text{ is orthonormal} \right\}.
\]

In 1981, Westwick proved that this set \( W_C(A) \) was convex.

If \( C = I_1 \oplus O_{n-1} \), then \( W_C(A) = W(A) \).

If \( C = I_k \oplus O_{n-k} \), then the range \( W_C(A) = W_k(A) \) is said to be the \( k \)-numerical range of \( A \).

If \( C \) is a complex diagonal matrix, then the range \( W_C(A) \) is also defined as the above. But it is known that if \( A = C = \text{diag}(1, (-1 + \sqrt{3}i)/2, (-1 - \sqrt{3}i)/2) \), then the range is the
closed region bounded by the deltoid

\[ \{2 \exp(i\theta) + \exp(-2i\theta) : 0 \leq \theta \leq 2\pi\}, \]

and hence it is not convex. For arbitrary \( n \times n \) complex matrices \( A, C \),
the range \( W_C(A) \) is defined as

\[ W_C(A) = \{ \text{tr}(CUAU^{-1}) : U \in \mathbf{M}_n, UU^* = I_n \}. \]

If \( C \) is a rank one matrix

\[
C = \begin{pmatrix}
q & \sqrt{1-q^2} & 0 & \ldots & 0 \\
0 & 0 & 0 & \ldots & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & 0 & \ldots & 0
\end{pmatrix},
\]

\(( -1 \leq q \leq 1)\), then the range \( W_C(A) \) is denoted by \( W(A : q) \).
Tsing [27;1984] proved that $W(A : q)$ was convex. We find that

$$W(A : q) = \{ \eta^* A \xi : \xi, \eta \in \mathbb{C}^n, \xi^* \xi = \eta^* \eta = 1, \eta^* \xi = q \}.$$ 

The $q$-numerical ranges satisfy $W(A : -q) = -W(A : q)$ and

$$W(A : 0) = \{ z \in \mathbb{C} : |z| \leq ||A|| \},$$


So the family of the $q$-numerical ranges $\{W(A : q) : 0 \leq q \leq 1\}$ interpolate the numerical range $W(A)$ and the circular disc $W(A : 0)$ with radius $||A||$.

Here I present an example of $q$-numerical ranges of a matrix. Suppose that

$$A = \begin{pmatrix} 0 & 1 & 1/2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad q = \frac{13}{14}. $$
Fig. 10 the numerical approximation of the range $W(A : 13/14)$ by using C. K. Li's computer program.
5. (2) application to NMR

Recently S. J. Glaser, U. Helmke, T. Schulte-Herbruggen et al [10;1998] applied the $C$-numerical range to the NMR (Nuclear Magnetic Resonance)-spectroscopy (cf. [13;2002]).

Fig.11 MRI of a brain
6. Kippenhahn’s method to compute the boundary

Let $A \in M_n$. The range $W(A)$ satisfies the equation

$$\text{max}\{\Re(z \exp(-i\theta)) : z \in W(A)\}$$

$$= \text{max} \sigma([\exp(-i\theta)A + \exp(i\theta)A^*]/2)$$

for $0 \leq \theta \leq 2\pi$. By using this equation, we know the equations of support lines of the compact convex set $W(A)$. Usually support lines are tangents of the curve $\partial W(A) \subset \mathbb{C} \cong \mathbb{R}^2$. We shall consider the curve

$$\Gamma = \{(X, Y) \in \mathbb{R}^2 : Xx + Yy + 1 = 0 \text{ is a tangent of } \partial W(A),$$

$$\text{in } (x, y) - \text{plane}\},$$

where we use the coordinates $x + iy$ on the Gaussian plane $\mathbb{C}$. This curve lies on an algebraic curve:

$$\tilde{\Gamma} = \{(X, Y) \in \mathbb{R}^2 : \det(I_n + X\Re(A) + Y\Im(A)) = 0\},$$
where \( \Re(A) = (A + A^*)/2 \), \( \Im(A) = (A - A^*)/(2i) \). For every boundary point \( x + iy \) of \( W(A) \), the straight line \( xX + yY + 1 = 0 \) is a tangent line of \( \tilde{\Gamma} \) or a support of the convex domain surrounded by some part of \( \tilde{\Gamma} \).

The (real part of the) dual curve \( f(x, y) = 0 \) of \( \tilde{\Gamma} \) is called the Kippenhahn curve, or the boundary generating curve for \( W(A) \) (cf. [15;1951]).

By using the above property, we can produce a real polynomial \( f(x, y) \) of degree \( \leq n(n - 1) \) satisfying \( f(x, y) = 0 \) for every boundary point \( x + iy \) of \( W(A) \). The polynomial \( f(x, y) \) appears as the resultant of the polynomial

\[
G(x, y : X) = y^n \det(I_n + X \Re(A) + (-1/y - xX/y)\Im(A)) = \det(yI_n + yX \Re(A) + (-1 - xX)\Im(A))
\]

and its derivative

\[
G_X(x, y : X) = \frac{\partial}{26\partial X} G(x, y : X)
\]
with respect to $X$, that is, we take the set of points $(x, y)$ for which the equation $G(x, y : X) = 0$ in $X$ has a repeated root.

We present an example of Kippenhan curves.

Suppose that

$$A = \begin{pmatrix} 1 & 1/5 & 3/5 \\ 0 & 1 & 2/5 \\ 0 & 0 & 0 \end{pmatrix}.$$ 

The figure in the next page shows the Kippenhahn curve for $W(A)$.

We call that the **discriminant** $\text{Dis}(g)$ of a polynomial

$$g(X) = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0 \quad (a_n \neq 0)$$

is the resultant of $g(X)$ and its derivative

$$g'(X) = na_nX^{n-1} + (n - 1)a_{n-1}X^{n-2} + \cdots + a_1$$ divided by $a_n$. 

27
Fig. 12  Kippenhahn curve for $W(A)$
Recently I have some interests in ancient astronomy and philosophy. A Greek philosopher Zeno of Elea (about 490 B.C -about 430 B.C):
paradoxes on motions

Aristotle: "Physica"

By virtue of Islamic scholars’s study on the ancient philosophy and science in Greece, it is known for the people in the world. For instance a Persian philosopher Ibn Sina (Abu, Ali, Avicenna, 980-1037) is famous for his philosophy and medical theory.

Zeno’s paradox concerns with mathematical contradictions, infinity, continuity, philosophical dialectics and ontology, etc.
Fig. 14  Achilles and a turtle
Classical mechanics and Astronomy is one of my recent interests. Ptolemaic theory provides epitrochoid models of geocentric planetary motion.
orbits. Johannes Kepler (1571-1630) supposed many different oval curve for Mars’s orbit before he found that the ellipse was the true orbit. I studied his oval orbits. One is quartic (cf. [14; 2008]) and another is octic. Astronomic theories before Copernicus was so complicated. A Persian astronomer, Nasir al-Din al-Tusi (1201-1274) is famous for his ”Tusi-couple” (cf. [29; Whiteside]). I feel some similarity between the Ptolemaic composition of movements and the Minkowski sum of elliptical discs.
Suppose that $L_1, L_2, \ldots, L_k$ are elliptical discs with center at 0. Then the convex set

$$L_1 + L_2 + \cdots + L_k = \{z_1 + z_2 + \cdots + z_k : z_j \in L_j (j = 1, 2, \ldots)\}$$

is said to be the \textit{Minkowski sum} of $L_1, L_2, \ldots, L_k$. If $L_j = W(A_j)$ for some $2 \times 2$ matrix $A_j$ with $\text{tr}(A_j) = 0$, then the range

$$W_k(A) = \left\{ \sum_{j=1}^{k} \xi_j^* A \xi_j : \{\xi_1, \xi_2, \ldots, \xi_k\} \text{ is orthonormal} \right\}$$

satisfies $W_k(A) = L_1 + L_2 + \cdots + L_k$ where $A$ is the $2k \times 2k$ matrix defined as

$$A = A_1 \oplus A_2 \oplus \cdots \oplus A_k.$$
[My conjecture] Does the boundary of this range $W_k(\mathcal{A})$ lies on an algebraic curve of degree $k \times 2^k$?

It is also known that the equation

$$W(T \otimes I_m + I_n \otimes S) = W(T) + W(S')$$

holds for an arbitrary $n \times n$ matrix $T$ and an $m \times m$ matrix $S$.

![Minkowski sum of two ellipses](image)

Fig.16 Minkowski sum of two ellipses
8. Joint numerical range and its convex hull

Suppose that \( \{H_1, H_2, \ldots, H_m\} \) is an \( m \)-ple of \( n \times n \) Hermitian matrices. The joint numerical range \( W(H_1, H_2, \ldots, H_m) \) is defined as

\[
W(H_1, H_2, \ldots, H_n) = \left\{ (\xi^* H_1 \xi, \xi^* H_2 \xi, \ldots, \xi^* H_m \xi) \in \mathbb{R}^m : \xi \in \mathbb{C}^n, \xi^* \xi = 1 \right\}.
\]

This closed set is not necessarily convex. In the case \( n \geq 3, m = 3 \), the range \( W(H_1, H_2, H_3) \) is convex (Au-Yeung, Tsing [1;1983]). The convex hull \( \Omega \) of the range \( W(H_1, H_2, \ldots, H_m) \) satisfies
\[ \Omega = \{(x_1, x_2, \ldots, x_m) \in \mathbb{R}^m : x_1 y_1 + x_2 y_2 + \cdots + x_m y_m \leq g(y_1, y_2, \ldots, y_m) \text{ for } \forall \text{ unit vector } (y_1, y_2, \ldots, y_m) \in \mathbb{R}^m \} \]

where

\[
 g(y_1, y_2, \ldots, y_m) = \max\{x_1 y_1 + x_2 y_2 + \cdots + x_m y_m : (x_1, x_2, \ldots, x_m) \in W(H_1, H_2, \ldots, H_m)\} \\
= \max\{\xi^* (y_1 H_1 + y_2 H_2 + \cdots + y_m H_m) \xi : \xi \in \mathbb{C}^n, \xi^* \xi = 1\} \\
= \max\sigma(y_1 H_1 + y_2 H_2 + \cdots + y_m H_m).
\]

Hence this set \( \Omega \) contains the point \((0, 0, \ldots, 0)\) if and only if the largest eigenvalue of \(y_1 H_1 + y_2 H_2 + \cdots + y_m H_m\) is non-negative for every non-zero vector \((y_1, y_2, \ldots, y_m) \in \mathbb{R}^m\).
If $m$ is even and $H_{2j-1} = (A_j + A_j^*)/2$, $H_{2j} = (A_j - A_j^*)/(2i)$ ($j = 1, 2, \ldots, m/2$), then the range $W(H_1, \ldots, H_m)$ is denoted by $W(A_1, \ldots, A_{m/2})$.

Fig. 17
9. Polynomial numerical hulls

Let $A \in \mathbf{M}_n$. In Section 2, we showed

$$W(A) = V^1(A) = \{ \lambda \in \mathbb{C} : |\lambda + a| \leq ||A + aI_n|| \text{ for } \forall a \in \mathbb{C} \}.$$  

For a positive integer $k$, this object is generalized as

$$V^k(A) = \{ \lambda \in \mathbb{C} : |a_k \lambda^k + a_{k-1} \lambda^{k-1} + \cdots + a_0|$$

$$\leq ||a_k A^k + a_{k-1} A^{k-1} + \cdots + a_0 I_n||$$

for $\forall a_k, a_{k-1}, \ldots, a_0 \in \mathbb{C}$, or

$$V^k(A) = \{ \lambda \in \mathbb{C} : a_k \lambda^k + a_{k-1} \lambda^{k-1} + \cdots + a_0$$

$$\in W(a_k A^k + a_{k-1} A^{k-1} + \cdots + a_0 I)$$

for $\forall a_k, a_{k-1}, \ldots, a_0 \in \mathbb{C}$.  

39
A more crucial characterization was found by V. Faber et al. (1996) and by Anne Greenbaum (2002) [11]:

\[ V^k(A) = \{ \lambda \in \mathbb{C} : (0, 0, \ldots, 0) \in \text{Conv}(W(A - \lambda I_n, (A - \lambda I_n)^2, \ldots, (A - \lambda_n)^k))\}. \]

Thus it also satisfies

\[ V^k(A) = \{ \lambda \in \mathbb{C} : \min_{y_1^2 + \cdots + y_{2k}^2 = 1} \max \sigma(\Re(A - \lambda I_n) + y_2 \Im(A - \lambda I_n) + \ldots + y_{2k-1} \Re((A - \lambda I_n)^{k}) + y_{2k} \Im((A - \lambda I_n)^{k})) \geq 0 \}. \]

Using this equation, we see that \( \partial V^k(A) \) lies on an algebraic curve. In the case \( k = n \), let

\[ p(\lambda) = \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_0 \]
be the characteristic polynomial of $A$. Then we have

$$V^n(A) \subset \{ \lambda \in \mathbb{C} : p(\lambda) = 0 \},$$

and hence $V^n(A) \subset \sigma(A)$. If $\lambda$ is an eigenvalue of $A$ and $\xi \in \mathbb{C}^n$ is a unit eigenvector of $A$ corresponding to $\lambda$, then the equation

$$\xi^* q(A) \xi = q(\lambda)$$

holds for an arbitrary polynomial $q(z)$. Hence the inequality

$$|q(\lambda)| \leq \|q(A)\|$$

holds for an arbitrary polynomial $q$. Thus the equation $V^n(A) = \sigma(A)$ holds. The numerical range $W(A) = V^1(A)$ of a matrix $A$ is used to locate the eigenvalues of $A$.

For a matrix polynomial $Q(\lambda)$, the range $V^k(Q)$ is defined as

$$V^k(Q) = \{ z \in \mathbb{C} : 0 \in V^k(Q(z)) \}.$$
Some interesting results of $V^k(Q)$ has been found by A. Salemi and Gh. R. Aghamollaei (cf. [8], [9], [25]).

For instance, they gave the following result. Suppose that $Q(\lambda)$ is a normal matrix polynomial, that is, $Q(\mu)^*Q(\mu) = Q(\mu)Q(\mu)^*$ for any $\mu \in \mathbb{C}$. Then

$$\partial W(Q) \cap V^2(Q) \subset \sigma(Q).$$


10. Conclusion

In 1918 and 1919, German mathematicians Toeplitz and Hausdorff gave a foundation of the theory of numerical range of linear operators on a Hilbert space. In 90 years, the theory of the numerical range has become an active branch of the functional analysis and the numerical analysis. Iranians, Americans, Canadians, Spanish, Chinese, Japanese, etc, mathematicians in many different countries have contributed to the development of the theory of numerical ranges.

I hope that a great progress will be performed in many branches of mathematics under the international peace.

Thanks you for your attention.

Thank you again for the invitation to this conference.
References


(electronic).


